



HAL
open science

Romani and Contact Linguistics

Yaron Matras, Evangelia Adamou

► **To cite this version:**

Yaron Matras, Evangelia Adamou. Romani and Contact Linguistics. Yaron Matras; Anton Tenser. The Palgrave Handbook of Romani Language and Linguistics, Springer International Publishing, pp.329-352, 2020, 978-3-030-28104-5. 10.1007/978-3-030-28105-2_11 . halshs-02965202

HAL Id: halshs-02965202

<https://halshs.archives-ouvertes.fr/halshs-02965202>

Submitted on 13 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1. Introduction

Romani is a language that is permanently in contact: There isn't a single monolingual Romani-speaking community, and probably no adult individual is monolingual in Romani. Bilingualism almost always starts at an early age, often in the family. As a minority community, often socially marginalised and dependent on contacts with the majority population for their livelihood, Roma always acquire the language of the surrounding majority population, and often that of regional minority groups as well, in a situation where bilingualism is unidirectional; that is, Roma acquire other languages but members of neighbouring populations only very seldom acquire Romani. This sociolinguistic asymmetry in power relations makes Romani the recipient of contact influences from neighbouring languages. Since Romani populations are geographically dispersed but show very similar sociolinguistic relations with their respective neighbouring populations, Romani dialects exhibit a range of contact influences from different languages. All this makes Romani a fascinating 'laboratory' for the study of bilingualism and contact-induced language change. Not only does Romani offer an opportunity to compare the impact of diverse contact languages on a rather homogeneous stock of inherited structures, the historical migrations of Romani populations have led to changing contact constellations, resulting in successive layers of contact influences within individual dialects. This enables us to correlate the historical depth of contact with patterns of contact-induced change in particular areas of structure (see Matras 2002, chapter 8; and see Elšík and Matras 2006).

Our aim in this chapter is to review the contribution that the study of contact-induced change in Romani has had to developing concepts, methods and theory in contact linguistics, as well as to present a bird's eye view of structural contact phenomena that are assumed to have affected Romani as a whole and thus to be constitutive of the language, and those that characterise contact-induced changes in contemporary Romani dialects. We draw on descriptive work in Romani linguistics as well as on a large body of studies that have been

devoted specifically to contact phenomena in Romani. We also refer in some instances to data from the online Romani Morpho-Syntax Database (RMS). Not only is this the largest and most comprehensive online resource that documents the dialects of Romani (and indeed the dialects of any language, as far as we are aware), it is also the only descriptive-typological language documentation resource that systematically tags items (including word forms and grammatical morphemes, and in some instances entire categories) for borrowing and historical depth of borrowing (see below). As a language documentation resource, RMS itself thus demonstrates important advances in the study of contact linguistics.

2. Approaches to contact in Romani linguistics

2.1 Topics

Contact influences have figured prominently in the dialectological study of Romani since its early beginnings. Miklosich's (1872-1880) seminal work on the dialects of Romani gave much consideration to loanwords and used them to trace the migrations of Romani speaking population across Europe, and to propose a classification of the dialects of Romani based on the constitutive impact of the respective contact languages on individual dialect groups. On the basis of the shared Greek component in all Romani dialects, Miklosich asserted that a Greek-speaking area had been the "European homeland of the Gypsies" and the point of departure for subsequent migrations across the European continent.

The Indo-Aryan genealogical heritage of Romani had been well established since the appearance of Pott's (1844-1845) work. Nonetheless, many early discussions continued to group together the Indo-Aryan dialects of Romani with speech variants attributed to various itinerate populations. Of the latter, those that were characterized by a Romani-derived lexicon received attention as distinct contact varieties. Haugen (1949) described the use of Romani vocabulary embedded into Norwegian as an identity marker, questioning the 'linguageness' of such stylistic hybrids, while Hancock (1970) and Boretzky (1985) both drew parallels between such so-called 'mixed' varieties (vernacular forms of the majority language with a Romani lexical component) and creoles as in-group vernaculars that owed their emergence to a blend of structures of different origins. Thomason and Kaufman (1988) discussed Angloromani (the 'mixed' variety of English Romani Gypsies, which embeds Romani

vocabulary into vernacular English), based on constructed examples from the literature, as one of the examples of what they called ‘broken transmission’ of languages across generations. They proposed that Angloromani was a ‘mixed language’ that had emerged as a result of a wholesale replacement of Romani grammar by that of the contact language, English, a scenario that is disputed (see Matras 2010). The presence of Romani-derived core lexicon along with, in some cases, traces of grammatical formation (especially in word derivation), in several different attested varieties in itinerant communities led to the emergence of a comparative discussion context and the postulation of a language type now commonly referred to as ‘Para-Romani’ (see Bakker and Courtiade 1991; and see Bakker, this volume). From a purely structural viewpoint, such varieties have been defined as blends of Romani lexicon with non-Romani grammar, which emerged following the abandonment of Romani as an everyday, all-purpose community language, coupled with the motivation to flag group identity as well as to have a group-internal means of communication. Usage-based studies have since defined this as an ‘emotive mode’ of speech whereby scattered group-internal, usually Romani-derived lexical items are used to evoke solidarity and the processing of a turn, utterance, or discourse portion from the perspective of a particular set of shared values (see Matras et al. 2007; Matras 2010).

Not unrelated is the contribution of Romani vocabulary to other languages, in particular slang varieties. Lists of slang items of Romani etymology have been provided in both early and more recent studies (for an overview see Matras 2002, p. 249), with a number of contributions addressing theoretical implications including the association of Romani with anti-establishment defiance which makes it a prestige variety in the social periphery (cf. Leschber 1995; Matras 1998a; see also contributions to Matras 1998b).

The wealth of contact phenomena attested in Romani prompted a series of works aimed at compiling a comprehensive overview of borrowed structures across Romani dialects for particular categories (cf. Friedman 1985; Boretzky and Igla 1991, 1993; Matras 1994b, 1996; Elšík 2001). Drawing on the RMS Database, Elšík and Matras (2006) postulate borrowing hierarchies for Romani dialects, while Matras (1998b, 2009, 2015) draws on Romani examples to propose generalisations and universal predictions on the borrowability of structural categories. Modern descriptive discussions of Romani (e.g. Igla 1989; Matras 1994) generally give prominent attention to contact phenomena, and many case studies are devoted

to contact phenomena in individual dialects, such as Elšik's (2009) discussion of lexical loans in Selice Romani of Slovakia.

Surprisingly little attention has been devoted to sociolinguistic and discourse related aspects of contact in Romani. In what is a rather isolated case study, Réger (1979) discusses second language acquisition, while Matras (2009) draws on Romani examples in his discussion of codeswitching. Halwachs (1993) introduces patterns of language maintenance and language shifts in the multilingual repertoires of Romani migrants, while Adamou and Granqvist (2015) and Adamou (2016) introduce a corpus-based evaluation of Romani switching patterns with a focus on inter-speaker variation. Leggio (2015) discusses the use of multilingual language resources in a Romani online forum, Leggio and Matras (2017) discuss the impact of contact on the use of orthographic variants in social media, and Abercrombie (2018) shows how a multilingual repertoire shapes attitudes to language and standardisation.

2.2 Concepts

Alongside the many descriptive contributions, discussions of contact phenomena in Romani have given rise to a number of novel concepts that have enriched the field of contact linguistics more generally. All Romani dialects display a split in the assignment of lexicon (nouns, verbs and adjectives) to two distinct types of inflection, the first continuing early Modern Indo-Aryan (with traces of late Middle Indo-Aryan) while the second relies on inflectional markers adopted from Byzantine Greek. Generally, pre-European lexicon is assigned to the first, while European loans, including most (but not all) Greek loans, are assigned to the second, which is thus the more productive for new vocabulary except for internal derivations. This split had been coined, using a term borrowed from Indo-European linguistics, 'thematic' (i.e. pre-European) vs. 'athematic' (i.e. European) grammar, by Kaufman (1979), though the terminology is often attributed to Hancock (1995 and earlier manuscripts). Bakker (1997) demonstrates that Greek-derived so-called 'athematic' morphology replicates a pattern used in Greek to adopt loanwords from other languages, while Elšik (2000) introduces etymology systematically as one of the dimensions of the Romani nominal paradigm, adopting in later work the more bespoke terms 'oikoclitic' (indigenous, i.e. pre-European) and 'xenoclitic' (used for loans, i.e. European) (cf. Elšik and Matras 2006; see also Elšik, this volume).

Controversial is the notion of Romani as a ‘deficient language’, proposed by Boretzky (1989) in an early attempt to provide an overview of contact behaviour in the language. Boretzky’s claim that the wholesale borrowing of entire categories into Romani (notably in the domain of particles, especially adversatives) might point to the absence of those structural categories in the pre-European forerunner or ‘Proto-Romani’ is refuted by Matras (1998), who links the wholesale adoption of certain categories to their discourse functionality as markers of ‘monitoring-and-directing’ of hearer participation and therefore the bearers of a cognitive load linked to the speaker’s degree of assertive authority when dealing with propositions that challenge or are not directly inferable from a shared presupposition domain. Matras (1996, 1998) introduces the term ‘fusion’ to account for such wholesale replacement and demonstrates that the process observed in Romani can also be found in other languages in contact. Hinting at that observation, Auer (1999) introduces the term ‘fused lect’ with reference to examples taken from Holzinger’s (1993) description of Sinti Romani. Opinions have since been split, with some accepting Auer’s position that the high rate of German lexical and grammatical borrowings appearing alongside spontaneous lexical insertions from German in the speech of the Romani-German bilingual Sinti Romani community offers an example of an in-between stage that could link ‘borrowing’ in the plain sense and the formation of a ‘mixed language’. Others regard Sinti Romani, much like all other Romani dialects and indeed countless minority or indigenous languages in contact with a powerful majority language, as a case of far-reaching borrowing which nevertheless demonstrates notable constraints and does not, in fact, offer any pathway for a gradual transition to the kind of languages that have been labelled ‘mixed languages’ (as defined in Bakker and Mous 1994; Matras and Bakker 2003, and elsewhere) and which combine predication grammar from one source with core lexicon and/or nominal grammar from another.

Above we already mentioned the term ‘Para-Romani’ (Bakker and Courtiade 1991), widely used to describe the replication of a sizeable structural component, primarily lexical and some grammatical relics, following language shift from Romani to the majority or regional language. For some authors (cf. Matras 2010) the concept entails a notion of post-vernacular language use that is associated on the one hand with language death and so-called ‘afterlife’, and on the other hand with the need to maintain a distinct in-group mode of communication akin to the functionality of secret lexicons, widely found among various itinerant populations, whether of Romani or other background. Other authors have argued for a distinction between

Para-Romani varieties and secret languages (cf. Bakker 1998), regarding the former as part of a class of ‘intertwined’ languages that owe their emergence to a pre-determined combination of lexicon and grammar from distinct sources, as part of the process of new identity-building. Other explanations offered for Romani ‘mixed dialects’ (Boretzky and Igla 1994) suggest a more volatile process of identity re-negotiation, possibly as part of a process of partial assimilation into indigenous itinerant populations (cf. Hancock 1992). Noteworthy is thus the widespread adoption in Romani linguistics of the notion of a ‘Para-’ language as one that is historically related but functionally and structurally confined. With a focus on this constrained and specialised functionality, some authors have associated Para-Romani with wholesale euphemism/dysphemism (Burridge and Allen 1998) or bystander-oriented deixis (Rijkhoff 1998), and more recently with an ‘emotive mode’ of communication (Matras et al. 2007; Matras 2010).

Shifting contact languages in the history of Romani as a whole and its individual varieties have inspired a model that distinguishes historical layers of contact. First introduced by Matras (1998) and then adopted as a category in the RMS Database and subsequently in discussions of borrowing hierarchies in Romani (Matras 2002; Elšik and Matras 2006), the model identifies, for each variety of the language, an Older, Recent and Current contact language (or languages). The Older L2 is one that has had considerable, prolonged impact on the forerunner of a particular dialect. Speakers, especially elderly speakers, are often aware of this impact, even if that L2 is no longer spoken by members of the community. The Recent L2 is the contact language that is no longer used by the entire community of speakers of a particular Romani dialect as their everyday language outside the home, but which may still be used by the parent or grandparent generation (or by the first generation of immigrants, in migrant communities), and to which the younger generation may still be exposed, at least occasionally. The current L2 is the principal contact language used by the community for everyday interaction with the non-Romani majority, and often as a family language alongside Romani. For example, Lovari Romani as spoken in Germany (Maras 1994a), has Romanian as Older L2, Polish as Recent L2, and German as Current L2, while Ajia Varvara Vlach (Igla 1989) has Romanian as Older L2, Turkish as Recent L2, and Greek as Current L2. Contact languages change, however, not only as a consequence of the migration of Roma, but sometimes as a result of shifting power relations among languages in situ. Thus, in some areas of southern Slovakia, Hungarian has shifted from Current to Recent L2 as Romani speaking communities adopted Slovak as Current L2. The method of stratifying contact influences in

diachronic depth has made it possible to assess the susceptibility of individual structural categories to borrowing and often to repeated, ‘replacive’ borrowing and convergence, laying the foundations for a model of implicational hierarchies of grammatical borrowing (Matras 1998, 2009) as well as re-positioning the discussion of borrowing within a model of markedness (Elšik and Matras 2006).

In regard to the latter, Elšik and Matras (2006) describe borrowing, in its incipience, as a strategy that supports bilingual speakers in successfully managing language choices in interaction: It reduces the need for choices to be made among word-forms or morphemes of equivalent function, and so it increases communicative efficiency without compromising the separation of languages and language contexts and the potential for flagging identity that this separation entails (see also Matras 2009: 159-160). For that reason, there is no single correlation between borrowing and ‘markedness’, as defined by criteria such as complexity and internal differentiation. Discussing implicational borrowing hierarchies as evidenced by a sample of Romani dialects, Elšik and Matras (2006: 371) show how different hierarchies align themselves in different ways with respect to the features ‘complexity’ and ‘differentiation’, respectively. For example, the value ‘plural’ is more prone to borrowing than the value ‘singular’, and is at the same time ‘marked’ as showing higher complexity and lower differentiation than singularity. By contrast, the value ‘nominative’ (in nominal case paradigms) shows higher borrowability than the value ‘oblique’, but is unmarked through lower complexity and greater internal differentiation.

Work on Romani has also had an impact on our understanding of convergent processes, or ‘pattern replication’ (Matras 2009). Friedman’s (1985) discussion of particular structures of Romani in the Balkan context was followed by Matras’s (1994b) attribution of wholesale syntactic-typological convergence in the Balkans to Early Romani, coining the concept of a ‘Balkanised New Indo-Aryan language’. The concept of ‘contact-induced functionalisation’ of categories (Matras 1994a, 1998d) as applied to Romani played a role in inspiring Heine and Kuteva’s (2005) much cited model of contact-induced grammaticalisation. Case specific discussions on the overall theme of convergence have been devoted to the re-functionalisation of inherited subject clitics in Romani as a means of adopting typological distinctions from the contact language, such as German verb-subject inversion (Matras 1999) or the split among attributive predications in Spanish (Adamou 2013), as well as the functionalization of Romani location expressions to replicate German aktionsart particles (Iglá 1992), among others.

Tenser (2016) introduces the concept of ‘semantic map borrowing’ in connection with the re-functionalisation of case markers in Northeastern Romani dialects in contact with Russian, Polish and Latvian.

A number of Romani dialects show exceptional contact behaviour in their replication of verb conjugation paradigms with borrowed lexical verbs. Iglá (1989) first pointed out the survival of Turkish paradigms with Turkish lexical verbs in the Ajia Varvara dialects of Greece. This phenomenon, since attested in various Romani dialects of the Balkans (and incipiently, in other regions as well) has been described as a form of ‘compartmentalized’ grammar (Friedman 2013; Matras 2015) or as a stage in an interrupted process of mixed language creation (Adamou 2010; Adamou and Granqvist 2015). Introducing methods from experimental psycholinguistics to the study of Romani, Adamou and Shen (2019) demonstrate that among Romani populations that also speak Turkish, language mixing is a dynamic process. Their experimental results show that cognitive cost in sentence processing depends on whether Turkish verbs have a more frequent Romani counterpart variant.

Finally, while the focus on contact behaviour in lexicon has tended to be on the rate of retention of pre-European (Indo-Aryan) words (cf. Boretzky 1992), viewed from the perspective of lexical borrowability Romani has since been presented as the reference point for ‘heavy borrowing’, showing a staggering borrowing rate of 62.7% on the comparative measure introduced by Haspelmath and Tadmor (2009) (cf. Elšik 2009). When comparing rates of contact words in natural speech (all word-classes considered), Adamou (2016) also finds that the Romani corpora are on the upper side of a language mixing scale, showing 20-35% tokens from the Current contact language(s).

3. Proto-Romani and Early Romani as languages in contact

The pre-European forerunner of Romani (‘Proto-Romani’) demonstrates a number of traits that point to convergent developments with the so-called Dardic or northwestern frontier languages of the Indian sub-continent into what will have been a Central Indo-Aryan language (see Matras 2002). In particular, the evidence is apparent in the presence of agglutinated case affixed and the renewal of the past tense conjugation. The transitivising suffix *-ar-* may have been a direct borrowing from the northwestern languages. Pre-European lexical loans point to

contact with Iranian languages and Armenian (see Scala, this volume), and it appears that some grammatical items, such as the adjectival comparative marker *-eder*, the adjectival prefix *bi-* ‘without’, the indefinite marker *či*, and the modal *šaj* ‘can’ may have been Iranian borrowings, while the nominal derivation marker in *-ik* is shared with Iranian and Armenian. Several syntactic-typological properties of Romani, such as the prepositioning of local relation adverbs (i.e. the development of prepositions), the emergence of external (‘remote’) tense markers that follow person affixes (*s-om-as* be-1SG-REM ‘I was’), the reduction of the infinitive in modal constructions, the use of resumptive pronouns with head nouns in non-subject roles in relative clauses, and the loss of Middle Indo Aryan relativisers in *y-/j-* and reliance instead on conjunctions derived from interrogatives, could have emerged in Proto-Romani in contact with Iranian languages.

Constitutive of ‘Early Romani’ – the term used to designate the common European or Anatolian ancestor of present-day Romani dialects for which individual forms can be reconstructed drawing directly on material that is continued in contemporary dialects – was contact with Byzantine Greek. Apart from several dozen lexical items of Greek origin that tend to be retained by Romani dialects today, as well as grammatical lexicon such as the adverbs *pale* ‘again’, *palpale/parpale/papale* ‘back’, and *komi* ‘still’, the word for ‘tomorrow’ (*tasja/tajsa/taha-*), and the numerals *efta* ‘7’, *oxto* ‘8’, *enja* ‘9’ as well as higher numerals, Greek has had a striking influence on Romani morphology, both derivational and inflectional. In derivation, we find the suffix for ordinal numerals *-to* (as in *duj-to* ‘second’), the nominaliser *-imo(s)* pl. *-imata*, indefinite *-moni*, participial *-imen* and the adjectival derivational suffixes *-itik-/itk-/ick-/ik-*, among others. In inflection, Romani adopts nominal class markers *-o(s)*, *-i(s)* and *-a*, plural *-i* from Greek, and adjectival endings *-o* and *-a*. With verbs, Greek tense/aspect endings, present *-iz-/az-/in-* etc. and aorist *-s-*, are adopted. These inflectional configurations constitute the ‘athematic’ or ‘xenoclitic’ grammar paradigms that serve to integrate loan vocabulary into subsequent contact languages in the individual dialects. With nouns, the Greek inflection template figures in the nominative case, and serves as a basis to which inherited non-nominative case markers are added (see Elšik, this volume). With verbs, the Greek aspectual markers serve as derivational morphemes that identify the verb stem as a loan, often in combination with inherited valency markers, and are followed by inherited person markers. Some Romani dialects maintain a Greek 3SG present-tense person marker *-i* with loan verbs.

It is plausible that many of the Balkan properties of Romani morpho-syntactic typology emerged in contact with Greek. This appears to be most obvious in the development of a pre-posed definite article from inherited deictic elements (cf. Matras 2002), as well as the grammaticalisation of oblique pronouns from demonstratives. The distinction between factual and non-factual clauses and complementizers (see Adamou and Matras, this volume) may also owe its history to convergence with Greek, like the shift to VO word order and the variable position of subject and verb in the sentence. The structure of adverbial and relative clauses tends to be compatible with Greek, though it may have its origin in pre-European (Western Asian) contacts.

4. Contact in contemporary Romani dialects

Detailed discussions of the impact of principal contact languages and language groups appear in this volume in the contributions by Meyer (on Slavic), Friedman (on Turkish), and Bodnarová and Wiedner (on Hungarian). Individual Romani dialects have absorbed influences also from most other European languages including Modern Greek, Romanian, Italian, French, Spanish, Basque, and Finnish. In the following we provide a brief sketch of typical influences in individual areas of structure.

4.1 Phonology

Although there is a lack of studies on this particular area, it is evident the prosody is particularly susceptible to contact influences and Romani dialects are often recognizable by their distinct prosody patterns which mirror elements of their current contact languages. Arvaniti and Adamou (2011) draw attention to the similar intonation patterns in *wh*- and polar questions as realized in Greek Thracian Romani and Modern Greek (see Adamou 2016 for more details and Grigorova 1998 for a similar study on Romani and Bulgarian prosodic convergence). Stress patterns remain conservative in many dialects, with accentuation falling on root or particular grammatical affixes (see Baló, this volume), but some Romani dialects of Western Europe have, by and large, shifted to word-initial stress, in contact with English and German, while the dialect of Prizren shows, in contact with Albanian, penultimate stress. Stress shifting patterns have been noted in individual dialects, where stress occasionally shifts to an earlier syllable than the one habitually carrying stress. The reasons for this phenomenon

are not entirely clear; they could be related to speaker's mixed dialect background, with focus marking replicating the contact language's model, or to metrical reasons (see Adamou and Arvaniti 2014). Northwestern and Central European dialects of Romani often adopt distinctive vowel length in contact with English, German, Finnish and Hungarian. The inventory of phonemes is, as a rule, enriched by the incorporation of lexicon from the current and often earlier contact languages, but borrowed phonemes are often restricted to borrowed lexicon. The phonemes /f/ and /ts/ are present primarily in loanwords but entered the language during the period of pre-European contacts. Loan phonemes from contemporary contact languages include rounded and central vowels, diphthongs, and palatals. There is a discernible adoption of phonetic values (points and mode of articulation) which tends to affect the inherited component and is thus subject to 'backwards diffusion' (Matras 2009). Examples are the articulation of /r/ as uvular [R] in contact with German, or the realization of /a/ as [ɑ] reported for British Romani. In many cases, such shifts in articulation may lead to changes in the phonemic system; such is the tendency toward merger of /s/ and /š/ in Greek dialects of Romani, the velarisation of /l/ to /ʎ/ in contact with Polish, the differentiation of /e/ and /ə/ in contact with Romanian, consonant gemination in contact with Finnish and Italian, the merger of the glottal /h/ with velar /x/ in contact with Russian and Greek, and the adoption of palatalization of dental and velar consonants in many Romani dialects of eastern Europe.

4.2 Morphology

Borrowed nominal derivational morphology from contemporary contact languages, including agentive, diminutive, and feminine markers, typically accompany borrowed nouns but do not always diffuse to the pre-European lexical component (but for examples of such diffusion, see for example Meyer, this volume). There is, however, a widespread tendency to rely on borrowed derivational markers for indefinite expressions. These include Romanian-derived *vare-*, *-godi*, *i-*, *bilo-*, *de-*, and *se-* (South Slavic), *-far* (Albanian), *vala-*, *akar-* (Hungarian), *nibud'* (Russian), and more. Romani dialects in contact with Polish, Russian, Ukrainian and Slovak, and to a lesser extent those in contact with Lithuanian, Bulgarian, and Greek, borrow productive aktionsart prefixes to modify inherited verb stems. Polish Romani, for example, shows almost wholesale borrowing of the full set of aktionsart (Slavic aspect) prefixes, leading to derivations like *za-pindžkirel* 'to introduce' from *pindžkirel* 'to recognize' (Polish *za-poznać* and *poznać*), *do-resel* 'to obtain' from *resel* 'to arrive' (Polish *do-stąpić* and

stapić), *pše-džal* ‘to cross, climb over’ from *džal* ‘to go’ (Polish *prze-chodzić* and *chodzić*). An attestation of productive borrowing of word-changing morphology is British Romani *ladj-fully* ‘shamefully’, while in Romani dialects in contact with Hungarian, the nominalising suffix *-išág-* is adopted (see Bodnárová and Wiedner, this volume, also for additional borrowed derivational affixes of Hungarian origin). Alongside the borrowing of morphological forms, Romani dialects in contact with strongly agglutinative languages, notably Hungarian and Turkish, show an increase in the productivity of inherited valency morphology (causative and passive) (see Elšík, this volume).

Borrowed inflectional morphology is relatively widespread in nominative plural markers, such as *-urj-/uri* of Romanian origin in Vlax, *-e* of Southern Slavic origin in the (predominantly western) Balkan dialects, and *-ides* of Greek origin in the eastern Balkan dialects. All these attach to borrowed European nouns, not just from the source language of the markers themselves, and Vlax *-urj-* even attaches to some Greek loans that entered the language prior to the period at which the suffix itself was acquired (e.g. *for-uri* ‘towns’, from Greek *foros*). It is also common in the inflection of adjectives to form comparative and superlative markers, where those are inflected morphemes in the contact languages, as in Hungarian (see Bodnárová and Wiedner, this volume). Both nominal plurals, and comparative and superlative markers on adjectives, can in fact be considered as borderline items between inflection (which relates to a constituent’s position in the sentence or discourse) and derivation (which affects standalone meaning).

A number of curious instances of morphological borrowing can be found across the dialects. The first is the borrowing of plural inflectional endings that attach to personal pronouns, where the shape of the pronouns in the contact language is similar, leading to an outcome where Romani adopted plural third person pronouns that are identical or near-identical to those of the contact language: The original Romani third person pronouns are *ov/oj* ‘he/she’, and *on* ‘they’; they appear in most Romani dialects, sometimes with minor phonological stem modifications. The form of the third person plural pronoun in some varieties of Hungarian Romani (Romungro) is *on-k*, which replicates the plural ending of the Hungarian third person pronoun (singular *ő*, plural *ők*). In Slovenian Romani, the form is *on-i*, replicating the pronominal form in Slovene (singular *on*, plural *on-i*), and in some varieties of Thracian Romani that are or were in contact with Turkish the form is *on-nar*, replicating the Turkish structure (singular nominative *o*, singular oblique *on-*, plural *on-lar*). In all three contact

languages, the ending that is used to indicate plurality on the pronoun is also the ending that is used to indicate plurality on nouns.

A somewhat comparable example is found in the past-tense verb conjugation paradigm of a number of Romani dialects of Bulgaria in historical contact with Turkish. The inherited Romani past-tense concord markers contain the consonant *-m* in the first person (singular *-om/-em/-im*, depending on dialect, plural *-am*) and a consonant *-n* in the second person (singular *-an*, plural *-en*). They resemble the corresponding Turkish singular forms 1SG *-Vm* and 2SG *-Vn* (with variation subject to vowel harmony). The Turkish plural pronouns are augmented forms of the singular morphemes: 1PL *-VmVz*, 2PL *-VnVz*. By analogy, these Romani dialects form a past-tense 1PL concord marker *-amus* and a past-tense 2PL marker *-enus*. Here too, the agglutinative marking of plurality in the contact language makes the marker *-us* analysable. It is replicated in Romani with inherited verbs, replacing the original marker (which is preserved in other dialects of the language). A similar process is found in Slovene Romani. Here, the original Romani past-tense 1PL marker *-am* has been replaced by the corresponding Slovene affix *-amo* on the basis of the formal resemblance between the two. The analogy is then extended to the Romani 2PL (originally *-an* or *-en*), for which the Slovene affix *-ate* is adopted (which has no formal resemblance to the original Romani form).

A ‘genuine’ borrowing of a pronominal form is found in Molisean Romani. Here, inherited object pronouns are cliticised, replicating the structure in the regional Italian vernacular. But for the 1PL and the reflexive pronoun, the Italian form *čə* is adopted: *dikkajom-čə* ‘I hurt myself’. English Romani, now extinct but documented in a series of amateur notes spanning two centuries (cf. Matras 2010), shows the borrowing of the English genitive *-s* in word compositions with inherited material, as in *daval’s tem* ‘heaven’ (literally ‘God’s country’), as well as the English gerundial marker *-in(g)*, as in *mandi sas wel’in keri* ‘I was comin’ home’. For the most part, borrowed inflectional endings on verb are confined to borrowed lexicon. As mentioned above, many Balkan Romani dialects show systematic compartmentalization whereby Turkish verb inflection accompanies Turkish derived lexical verbs:

Kalburdžu Romani dialect of Sindel in Northeastern Bulgaria (RMS, BG-008):

- | | | | | | | |
|-----|----------|-------------------|-------|----------------|------|----------|
| (1) | pandž-e | daka-en-da | sona | bašla-də | te | konušu-i |
| | five-OBL | minute-OBL.PL-ABL | after | begin-PAST.3SG | COMP | talk-3SG |

‘After five minutes he started to talk’

Here, in fact, the Turkish derived verb ‘begin’ takes the Turkish derived conjugation ending of the 3SG in the preterite, while the equally Turkish derived verb ‘talk’ in the subjunctive takes the Greek-derived 3SG ending *-i* that is reserved for loan verbs in various other Romani dialects. It is not uncommon for borrowed modal verbs to retain the inflectional ending of the source language. Particularly common are frozen impersonal forms in the 3SG, such as *može* ‘it is possible’, from South Slavic, used as an impersonal expression ‘can’, but occasional person inflected forms are also found, such as Serbian Romani *mora-m* ‘I must’, which retains the Serbian 1SG ending. The following example from the Romani dialect of Parakalamos in Greece shows three distinct layers of Greek-derived verb inflection: The impersonal form in the 3SG in *prepi* ‘must, it is necessary’, the person-inflected form *bor-o* ‘I can’, and the incipient use of the Greek person inflection in the Greek derived lexical verb *diavaz-o* ‘I read’, contrasting with the integration into Romani inflection of another Greek derived verb, *vojt-iz-av* ‘I help’ (drawing on the loan verb integration marker *-iz-* borrowed from Greek into Early Romani):

Parakalamos in the northwestern Greek province of Epirus (RMS, GR-002):

- (2) na *bor-o* te *diavaz-o* soske *prepi* te *vojt-iz-av*
NEG can-1SG COMP study-1SG because must COMP help-LOAN-1SG
me daj-a
my.OBL mother-OBL
‘I cannot study because I have to help my mother’

4.3 Syntax and morpho-syntactic convergence

Various areas of structure show examples of ‘pattern replication’ (Matras 2009) or restructuring of form-meaning alignment. In nominal typology, we find tendencies toward loss of definite and indefinite articles under the influence of Polish and Russian, and of loss of gender agreement and gender distinctions under the influence of Finnish and Hungarian. While word order in the noun phrase generally remains intact, dialects in the Balkans, especially those in contact with Romanian and Greek, show optional stylistic postpositioning of demonstratives and adjectives (see Adamou and Matras, this volume). In contact with

Russian, Romani shows some significant changes to case alignment (see Tenser 2016), as can be seen in the following example:

Russian Romani, Ekaterinburg (RMS, RUS-008):

(3) man na sys kxere
I.OBL NEG was.3SG at.home
I was not at home

cf. Russian:

menya ne bylo doma
I.ACC NEG was.3SG at.home

cf. Polish Romani (RMS, PL-003):

me na somys khere
I.NOM NEG was.1SG at.home

In the verb phrase, dialects in contact with German and Hungarian show a tendency to calque aktionsart particles drawing on location expressions, as in Sinti Romani (Germany) *kerau pre* ‘I open’ (lit. ‘I.make up), based on German *ich mache auf*. In contact with German, there is a tendency to formalize verb-subject inversion when the first sentence position is occupied by another constituent (see Matras 1999), while in contact with Hungarian there is a tendency for the copula to appear in final position in declarative clauses. Romani dialects in contact with Western Slavonic languages tend to bring forward the pronominal object to pre-verbal position:

Bergitka Roma, Krakow (RMS, PL-007)

(4) jov łes na dikhla
he.NOM him.OBL NEG see.PAST.3SG
‘He didn’t see him’

cf. Macedonian Arli, Skopje (RMS, MK-002)

ov na dikhlja ole
he.NOM NEG see.PAST.3SG him.OBL

In the Romani dialects of Central Europe and Ukraine, subject agreement in modal complement clauses is neutralized and the subordinate verb takes on a uniform inflection, derived in most cases either from the 3SG or 2SG, leading in effect to the emergence of a ‘new infinitive’ (see Adamou and Matras, this volume). Calquing also leads to the emergence of auxiliary verbs in some dialects, as in the emergence of a perfect auxiliary based on the verb ‘to have’ (itself derived from the inherited verb ‘to hold’) in contact with Greek – *therav kerdo* ‘I have done’ – or based on the verb ‘to be’ in dialects in contact with Macedonian, and of a progressing auxiliary from the verb ‘to be’ in English Romani: *shum to jaw* ‘I am going to walk’.

Adamou (2013) reports that Mexican Romani speakers developed a distinction between attributive predications using the copula *si* ‘to be’ and the third person subject clitic pronouns in *l-*. She argues that this innovation aims to replicate the Spanish *ser* vs. *estar* ‘to be’ copula variation, a claim that has since been confirmed by quantitative evidence (Padure, De Pascale, and Adamou 2018). This increase in complexity in Mexican Romani also offers an interesting counter-example to the hypothesis of simplification of alternatives among bilinguals as a strategy to reduce cognitive load. However, when Adamou et al. (2018) compare the Romani-Spanish copula preferences to those of Mexican Spanish monolinguals, they find that simplification of alternatives is currently taking place among bilinguals, with the generalization of the Romani clitics in third person affirmative clauses driving the generalization of the Spanish copula *estar*. The diachronic and synchronic Mexican Romani data therefore support bidirectional processes of conceptual transfer.

4.4. Grammatical vocabulary

As noted above, Romani has served as a model for the postulation of borrowing hierarchies in the domain of grammar and especially grammatical function words (Matras 1998; 2009), allowing a significant step in the revision of Thomason and Kaufman’s (1988) somewhat crude borrowing hierarchy in which ‘function words’ were treated as a wholesale, uniform category on the borrowing scale.

All Romani dialects borrow the adversative conjunction ‘but’ from the current or recent contact language, a testimony to its particular susceptibility to borrowing. The borrowing hierarchy ‘but’ > ‘or’ > ‘and’ was demonstrated for Romani dialects, with reference to the historical stratification of contact languages (older, recent and current L2). The same hierarchy has been identified in a number of other cross-linguistic samples, prompting the suggestion that long term borrowing is triggered by cognitive factors that relate to the processing of interrupted presuppositional chains, leading to ‘interference’ or spontaneous innovation and, if duly propagated, to structural change (Matras 1998). Romani discourse markers such as fillers and tags are invariably those of the current contact language, demonstrating the ‘fusion’ of so-called monitoring and directing operations across speakers’ repertoire of linguistic resources; this suggests that speakers have all linguistic resources available to them at all times, but also that Romani speech norms lack impediments on conformity to a distinctive set in these particular areas of structure, contrasting with what is otherwise an overall strong sense of language loyalty and cross-generation language transmission. Similarly, all phasal adverbs (‘still’, ‘already’, ‘no longer’), focus particles (‘even’, ‘only’, ‘every’, ‘also/too’), and sequential discourse markers (of the type ‘and then’, ‘and so’) are borrowed from recent or older contact languages.

Exclusively of European origin are the focal quantifiers ‘every’ (*svako/sako* from Slavic, *her/er* from Turkish), ‘entire/whole’ (*celo* from Slavic, *intrego* from Romanian, *-lauter* from German, *kre(j)t* from Albanian), ‘same’ (*isto* from South Slavic), and frequently also the ordinal ‘first’ (*pervo* from Slavic, *eršto* from German). In Balkan dialects in direct contact with Turkish, the Turkish numeral classifier *tane* is frequently replicated for the contrastive focus of a numeral (Adamou 2016). In the numeral system, ‘1000’ is usually a European loan: Sinti and Central dialects *ezero(s)* (Hungarian), Northeastern *tisač* (Slavic), Vlach *mija* (Romanian), Balkan *hilja* (Greek). Expressions for days of the week, time adverbs, and qualifying sentential adverbs (e.g. ‘probably’, ‘exactly’) also tend to be borrowed.

When it comes to other domains of function words, the picture is more mixed. Indefinite expressions are certainly among the categories with very high ‘matter’ replication, both of full indefinite expressions (such as, in respective dialects, Slavic *ništa* ‘something/nothing’, Romanian *uni* ‘some’, Hungarian *šoha* ‘never’, Greek *tipota* ‘nothing’, Turkish *hiç* ‘none’, and more), and of indefinite markers to which inherited ontological morphemes are added

(such as Vlax Romani *vare-so* ‘something, *vare-kon* ‘somebody’ etc., with *vare-* from Romanian).

Based on RMS data, we can also identify adpositions that are most prone to borrowing from contemporary (older, recent or current) contact languages:

- (5) Frequently borrowed adpositions by meaning:
opposite, between, against, among, around, through, without, instead of, except for, beyond, since, during, towards, with, from

As discussed by Elšík and Matras (2006, pp. 267-269), the general semantic hierarchy that emerges from this pattern is one in which peripheral relations (those involving more than one reference point, or movement) show greater borrowability than core relations. In addition, we encounter once again the semantics of contrast and discontinuity (and thus breaking of presuppositional inference) as a factor supporting borrowability (see Matras 2009). The final two positions on the list, ‘with’ and ‘from’, represent relations that are usually encoded in Romani by case relations but represented in individual dialects by borrowing of the preposition ‘with’ from German, Greek, Spanish or Italian, and of ‘from’ from German or Slavic. Borrowed adverbial subordinators are common especially for simultaneity (‘while’), cause and result (‘because’), concession (‘although’), and negative circumstance (‘without’), with a number of dialects borrowing conditional conjunctions, respectively, from Turkish, Finnish, Italian, Russian and Romanian. In contact with Slavic, Romani dialects tend to borrow the conditional particle that attaches to the verb phrase, *by*, *li* or *dali*. A number of interrogatives also show frequently borrowing, notably ‘how much’ and (less frequently) ‘when’.

The factual complementiser (inherited *kaj*) is replaced by loans from Hungarian, Italian, Greek, Bulgarian, (Balkan) Turkish, and Romanian in the respective dialects. Among modality expressions, ‘must’ shows high borrowability from recent or older contact languages (meaning that as a borrowing, it remains relatively stable and not susceptible to short term change of contact language). The potential modal ‘can’ is borrowed in some dialects, while ‘cannot’ is more rarely borrowed (see Elšík and Matras 2006, pp. 209-210).

4.5. Content lexicon

Surprisingly, for a language that is permanently in contact, and indeed has been described by some specialists as “deficient” (cf. Boretzky 1989) in that it is inherently dependent on material from neighbouring languages, very few studies have been devoted to lexical borrowing in Romani. As noted above, most studies on Romani lexicon tend to focus on the retention rate and etymological sources of pre-European vocabulary (cf. Scala, this volume). At the same time it is noteworthy that modern studies on lexical borrowing are generally few, and the field was given a boost only fairly recently through the comparative study of Haspelmath and Tadmor (2009). Part of that enterprise is Elšík’s (2009) work on a dialect of Slovak (Selice) Romani, which has become the benchmark for ‘heavy lexical borrowing’. Applying Haspelmath and Tadmor’s (2009) list of some 1,400 categorized meaning entries to Selice Romani, Elšík finds that borrowings make up over 90% of items in the domain ‘household, modern world, agriculture’, over 80% in the domain ‘clothing, warfare’, over 70% of words for ‘animals, social and political relations, and the physical world’, over 60% in ‘religion and belief, speech and language, law, technology, food and drink’, over 50% for ‘time, the body, motion, perception, emotion, cognition, values’, over 40% for ‘spatial relations’, and over 30% for ‘quantity, and kinship’. It is noteworthy, however, that ‘borrowing’ in the sense applied in the comparative study covers all items that are identified as having a non-core etymology, and thus for Romani it includes pre-European and Medieval Greek loans, and not just contemporary loans (older, recent, and current European contact language).

The RMS Database tends to confirm these findings, as far as the comparison allows, for Romani as a whole. It also allows us a more nuanced comparative examination among the dialects for individual expressions within the contact-susceptible categories. We find that in the RMS category ‘human beings’, frequent borrowings include the terms for ‘soldier’ and ‘friend’. For ‘animals’, the word ‘cat’ is usually borrowed. Among ‘body parts’, ‘fur’ and ‘skin’ are more likely to be borrowed, while in the domain ‘nature’ frequently borrowed items are ‘sea’, ‘weather’, ‘sky’, and ‘dust’. Frequently borrowed core verbs include ‘speak’, ‘pay’, and ‘love’.

Lexical semantic convergence is common, though this too is an area that has so far seen few if any dedicated studies. From the now extinct English Romani (cf. Matras 2010) we can mention collocations such as *ke divous* ‘today’, combining the Romani preposition *ke* ‘to’

with **dives* ‘day’, and expressions such as *sâr o čeros* ‘all the time’, where Romani **sa*, a plurality quantifier, shifts to become a determiner.

5 Conclusion

This chapter has outlined research and critical questions in contact linguistics involving Romani. In sum, it appears that language contact is to a great extent constitutive of Romani. In particular, Romani is associated to some cross-linguistically rare outcomes as it exhibits ‘heavy’ borrowing in both lexicon and grammar. Over the past decades, exploration of Romani dialects has given rise to new concepts, especially in the search for explanations for the emergence of mixed languages, either through substantial replication of paradigms leading to compartmentalization, or through near wholesale retention of lexicon after language shift.

An important goal for current and future research in Romani linguistics remains the understanding of abandonment of Romani and the transition to Para-Romani as well as of the continuum in borrowing and codeswitching as noted in some mixed dialects. In addition, comparison of contact behaviour in Romani and other languages can help refine borrowing hierarchies and offer an explanatory account (see Matras 2009), including by linking contact outcomes to their anchoring in language ecology (Adamou 2010). Lastly, we believe that the management of complex linguistic repertoires offers an avenue for future research in the study of Romani, with potential impact on the field of critical sociolinguistics. Such sociolinguistic and discourse-base studies are still missing in the context of Romani linguistics, and will hopefully become a priority for new researchers entering the field.

References

- Abercrombie, Amelia. 2018. Language purism and social hierarchies: Making a Romani standard in Prizren. *Language in Society*. 47(5):741-761.
- Adamou, Evangelia. 2010. Bilingual speech and language ecology in Greece: Romani and Pomak in contact with Turkish. *Language in Society*. 39(2):147-171.
- Adamou, Evangelia. 2013. Replicating Spanish *estar* in Mexican Romani. *Linguistics*. 51(6):1075-1105.

- Adamou, Evangelia. 2016. A corpus-driven approach to language contact. Endangered language in a comparative perspective. Berlin: De Gruyter Mouton.
- Adamou, Evangelia and Arvaniti, Amalia. 2014. Greek Thrace Xoraxane Romane. Journal of the International Phonetic Association. 44:223-231.
- Adamou, Evangelia, and Granqvist, Kimmo. 2015. Unevenly mixed Romani languages. International Journal of Bilingualism. 19(5):525-547.
- Adamou, Evangelia, De Pascale, Stefano, Garcia-Markina, Yekaterina, and Padure, Cristian. 2018. Do bilinguals generalize *estar* more than monolinguals and what is the role of conceptual transfer? International Journal of Bilingualism. First published on December 5, 2018. doi.org/10.1177/1367006918812175
- Adamou, Evangelia, and Shen, Rachel X. 2019. There are no language switching costs when codeswitching is frequent. International Journal of Bilingualism. 23(1):53-70.
- Arvaniti, Amalia, and Adamou, Evangelia. 2011. Focus expression in Romani. In: Byram Washburn, Mary, McKinney-Bock, Katherine, Varis, Erika, Sawyer, Ann, and Tomaszewicz, Barbara, eds. Proceedings of the 28th west coast conference on formal linguistics. Somerville, MA: Cascadilla Proceedings Project. 240-248.
- Auer, Peter. 1999. From codeswitching via language mixing to fused lects: toward a dynamic typology of bilingual speech. International Journal of Bilingualism. 3:309-332.
- Bakker, Peter. 1997. Athematic morphology in Romani: The borrowing of a borrowing pattern. In: Matras, Yaron, Bakker, Peter, and Kyuchukov, Hristo, eds. The typology and dialectology of Romani. Amsterdam: Benjamins. 1-21.
- Bakker, Peter. 1998. Para-Romani language versus secret languages: Differences in origin, structure, and use. In: Matras, Yaron, ed. The Romani Element in Non-Standard Speech. Wiesbaden: Harrassowitz. 69-96.
- Bakker, Peter, and Cortiade, Marcel. eds. 1991. In the margin of Romani. Gypsy languages in contact. Amsterdam: Institute for General Linguistics.
- Bakker, Peter and Mous, Maarten. eds. 1994. Mixed languages. 15 case studies in language intertwining. Amsterdam: IFOTT.
- Boretzky, Norbert. 1985. Sind Zigeunersprachen Kreols? In: Boretzky, Norbert, Enninger, Werner, and Stolz, Thomas. eds. Akten des 1 Essener Kolloquiums über Kreolsprachen und Sprachkontakte. Bochum: Verlag N. Brockmeyer. 43-70.
- Boretzky, Norbert. 1989. Zum Interferenzverhalten des Romani. (Verbreitete und ungewöhnliche Phänomene). Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung. 42:357-374.

- Boretzky, Norbert. 1992. Zum Erbwortschatz des Romani. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*. 45:227-251.
- Boretzky, Norbert, and Iglá, Birgit. 1991. Morphologische Entlehnung in den Romani-Dialekten. (Arbeitspapiere des Projektes "Prinzipien des Sprachwandels" 4). Essen: Universität GH Essen, Fachbereich Sprach- und Literaturwissenschaften.
- Boretzky, Norbert, and Iglá, Birgit. 1993. Lautwandel und Natürlichkeit. Kontaktbedingter und endogener Wandel im Romani. (Arbeitspapiere des Projekts "Prinzipien des Sprachwandels" 15). Essen: Universität GH Essen, Fachbereich Sprach- und Literaturwissenschaften.
- Boretzky, Norbert, and Iglá, Birgit. 1994. Romani mixed dialects. In: Bakker, Peter, and Mous, Maarten. eds. *Mixed languages. 15 case studies in language intertwining*. Amsterdam: IFOTT. 35-68.
- Bubeník, Vít. 1995. On typological changes and structural borrowing in the history of European Romani. In: Matras, Yaron. ed. *Romani in contact: The history, structure, and sociology of a language*. Amsterdam: Benjamins. 1-23.
- Burridge, Kate (with Keith Allan). 1998. The X-phemistic value of Romani in non-standard speech. In: Matras, Yaron. ed. *The Romani element in non-standard speech*. Wiesbaden: Harrassowitz. 29-49.
- Elšík, Viktor. 2000. Romani nominal paradigms: Their structure, diversity, and development. In: Elšík, Viktor, and Matras, Yaron. eds. *Grammatical relations in Romani: The noun phrase*. Amsterdam: Benjamins. 9-30.
- Elšík, Viktor. 2001. Word-form borrowing in indefinites: Romani evidence. *Sprachtypologie und Universalienforschung*. 54(2):126-147.
- Elšík, Viktor and Matras, Yaron. 2006. *Markedness and language change: The Romani sample*. Berlin: Mouton de Gruyter.
- Elšík, Viktor. 2009. Loanwords in Selice Romani, an Indo-Aryan language of Slovakia. In: Haspelmath, Martin and Tadmor, Uri. eds. *Loanwords in the world's languages. A comparative handbook*. Berlin: Mouton de Gruyter. 260-303.
- Friedman, Victor A. 1985. Balkan Romani modality and other Balkan languages. *Folia Slavica*. 7:381-389.
- Friedman, Victor A. 2013. Compartmentalized grammar: The variable (non)-integration of Turkish verbal conjugation in Romani dialects. *Romani Studies*. 23(1):107-120.
- Grigorova, Evelina. 1998. Interrogative intonation of two Bulgarian Romani dialects: Sofia Erli and Kalderaš. *Grazer Linguistische Studien*. 50:45-63.

- Halwachs, Dieter W. 1993. Polysystem, Repertoire und Identität. *Grazer Linguistische Studien*. 39-40:71-90.
- Hancock, Ian F. 1970. Is Anglo-Romanes a creole? *JGLS*, third series. 49:41-44.
- Hancock, Ian F. 1992. The social and linguistic development of Scandoromani. In: Jahr, Ernst Håkon, ed. *Language contact: theoretical and empirical studies*. Berlin: Mouton de Gruyter. 37-52.
- Hancock, Ian F. 1995. *A handbook of Vlax Romani*. Columbus: Slavica.
- Haspelmath, Martin and Tadmor, Uri. eds. *Loanwords in the world's languages. A comparative handbook*. Berlin: Mouton de Gruyter.
- Haugen, Einar. 1949. A note on the Romany "language". *Norsk Tidsskrift for Sprogvidenskap*. 7:388-391.
- Heine, Bernd and Kuteva, Tania. 2005. *Language Contact and grammatical change*. Cambridge: Cambridge University Press.
- Holzinger, Daniel. 1993. *Das Rómanes. Grammatik und Diskursanalyse der Sprache der Sinte*. Innsbruck: Verlag des Instituts für Sprachwissenschaft der Universität Innsbruck.
- Igla, Birgit. 1989. Kontakt-induzierte Sprachwandelphänomene im Romani von Ajia Varvara (Athen). In: Boretzky, Norbert, Enninger, Werner, and Stolz, Thomas. eds. *Vielfalt der Kontakte*. Bochum: Verlag N. Brockmeyer. 67-80.
- Igla, Birgit. 1992. Entlehnung und Lehnübersetzung deutscher Präfixverben im Sinti. In: Erfurt, Jürgen, Jeßing, Benedikt, and Perl, Matthias. eds. *Prinzipien des Sprachwandels I: Vorbereitung*. Bochum: Verlag N. Brockmeyer. 38-56.
- Kaufman, Terrence. 1979. Review of Weer Rajendra Rishi, *Multilingual Romani Dictionary*. *International Journal of the Sociology of Language*. 19:131-144.
- Leggio, D. Viktor. 2015. Radio Romani Mahala: Romani identities and languages in a virtual space. In Alietti, Alfredo, Olivera, Martin, and Riniolo, Veronica. *Virtual citizenship? Roma communities, inclusion policies, participation and ICT tools*. Milan: MacGraw-Hill Education. 97-114.
- Leggio, D. Viktor and Matras, Yaron. 2017. Orthography development on the Internet: Romani on YouTube. In: Jones, Mari C. and Mooney, Damien. eds. *Creating Orthographies for Endangered Languages*. Cambridge: Cambridge University Press. 254-275.
- Leschber, Corinna. 1995. Romani lexical items in colloquial Rumanian. In: Matras, Yaron. ed. *Romani in contact: the history and sociology of a language*. Amsterdam: Benjamins. 151-176.

- Matras, Yaron. 1994a. Untersuchungen zu Grammatik und Diskurs des Romanes. Dialekt der Kelderaša/Lovara. Wiesbaden: Harrassowitz.
- Matras, Yaron. 1994b. Structural Balkanisms in Romani. In: Reiter, Norbert, Hinrichs, Uwe, and van Leeuwen-Turnocová, Jefina. eds. Sprachlicher Standard und Substandard in Südosteuropa und Osteuropa. Wiesbaden: Harrassowitz. 195-210.
- Matras, Yaron. 1996. Prozedurale Fusion: Grammatische Interferenzschichten im Romanes. Sprachtypologie und Universalienforschung. 49:60-78.
- Matras, Yaron. 1998a. The Romani element in Jenisch and Rotwelsch. In: Matras, Yaron. ed. The Romani element in non-standard speech. Wiesbaden: Harrassowitz. 193-230
- Matras, Yaron. 1998b. Utterance modifiers and universals of grammatical borrowing. Linguistics. 36:281-331.
- Matras, Yaron. ed. 1998c. The Romani element in non-standard speech. Wiesbaden: Harrassowitz.
- Matras, Yaron. 1998d. Convergent development, grammaticalization, and the problem of 'mutual isomorphism'. In: Boeder, Winfried, Schroeder, Christoph, and Wagner, Karl-Heinz. eds. Sprache in Raum und Zeit. Tübingen: Narr. 89-103.
- Matras, Yaron. 1999. Subject clitics in Sinti. Acta Linguistica Academiae Scientiarum Hungaricae. 46:147-169.
- Matras, Yaron. ed. 1995. Romani in contact: the history and sociology of a language. Amsterdam: Benjamins.
- Matras, Yaron. 2002. Romani: A linguistic introduction. Cambridge: Cambridge University Press.
- Matras, Yaron. 2009. Language contact. Cambridge: Cambridge University Press
- Matras, Yaron. 2010. Romani in Britain: The afterlife of a language. Edinburgh: Edinburgh University Press.
- Matras, Yaron. 2015. Why is the borrowing of inflection morphology dispreferred? In: Gardani, Francesco, Arkadiev, Peter and Amiridze, Nino. eds. Borrowed morphology. Berlin: De Gruyter Mouton. 47-80.
- Matras, Yaron and Bakker, Peter. eds. 2003. The mixed language debate. Theoretical and empirical advances. Berlin: Mouton de Gruyter.
- Matras, Yaron, Gardner, Hazel, Jones, Charlotte, and Schulman, Veronica. 2007. Angloromani: A different kind of language? Anthropological Linguistics. 49(2):142-164.
- Miklosich, Franz. 1872-1880. Über die Mundarten und Wanderungen der Zigeuner Europas X-XII. Wien: Karl Gerold's Sohn.

- Padure, Cristian, De Pascale, Stefano, and Adamou, Evangelia. 2018. Variation between the copula *si* 'to be' and the *l*-clitics in Romani spoken in Mexico. *Romani Studies*. 28(2):263-292.
- Pott, August. 1844-1845. *Die Zigeuner in Europa und Asien. Ethnographisch-linguistische Untersuchung vornehmlich ihrer Herkunft und Sprache*. Halle: Heynemann.
- Réger, Zita. 1979. Bilingual Gypsy children in Hungary: Explorations in "Natural" second-language acquisition. *International Journal of the Sociology of Language*. 19:59-82.
- Rijkhof, Jan. 1998. Bystander deixis. In: Matras, Yaron. ed. *The Romani element in non-standard speech*. Wiesbaden: Harrassowitz. 51-67.
- Tenser, Anton. 2016. Semantic map borrowing - case representation in Northeastern Romani dialects. *Journal of Language Contact*. 9:211-245.
- Thomason, Sarah G. and Kaufman, Terrence. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.